MICROBIAL CELL FACTORIES

RESEARCH **Open Access**

Molecular cloning and biochemical characterization of a novel erythrose reductase from *Candida magnoliae* JH110

Dae-Hee Lee[1,3], Ye-Ji Lee[1], Yeon-Woo Ryu[2] and Jin-Ho Seo*[1]

Abstract

Background: Erythrose reductase (ER) catalyzes the final step of erythritol production, which is reducing erythrose to erythritol using NAD(P)H as a cofactor. ER has gained interest because of its importance in the production of erythritol, which has extremely low digestibility and approved safety for diabetics. Although ERs were purified and characterized from microbial sources, the entire primary structure and the corresponding DNA for ER still remain unknown in most of erythritol-producing yeasts. *Candida magnoliae* JH110 isolated from honeycombs produces a significant amount of erythritol, suggesting the presence of erythrose metabolizing enzymes. Here we provide the genetic sequence and functional characteristics of a novel NADPH-dependent ER from *C. magnoliae* JH110.

Results: The gene encoding a novel ER was isolated from an osmophilic yeast *C. magnoliae* JH110. The ER gene composed of 849 nucleotides encodes a polypeptide with a calculated molecular mass of 31.4 kDa. The deduced amino acid sequence of ER showed a high degree of similarity to other members of the aldo-keto reductase superfamily including three ER isozymes from *Trichosporonoides megachiliensis* SNG-42. The intact coding region of ER from *C. magnoliae* JH110 was cloned, functionally expressed in *Escherichia coli* using a combined approach of gene fusion and molecular chaperone co-expression, and subsequently purified to homogeneity. The enzyme displayed a temperature and pH optimum at 42°C and 5.5, respectively. Among various aldoses, the *C. magnoliae* JH110 ER showed high specific activity for reduction of erythrose to the corresponding alcohol, erythritol. To explore the molecular basis of the catalysis of erythrose reduction with NADPH, homology structural modeling was performed. The result suggested that NADPH binding partners are completely conserved in the *C. magnoliae* JH110 ER. Furthermore, NADPH interacts with the side chains Lys252, Thr255, and Arg258, which could account for the enzyme's absolute requirement of NADPH over NADH.

Conclusions: A novel ER enzyme and its corresponding gene were isolated from *C. magnoliae* JH110. The *C. magnoliae* JH110 ER with high activity and catalytic efficiency would be very useful for *in vitro* erythritol production and could be applied for the production of erythritol in other microorganisms, which do not produce erythritol.

Background

Erythritol, a four-carbon polyol, occurs naturally in a wide variety of foods including many fruits and mushrooms as well as in fermented foods such as cheese, wine, beer, and soy sauce [1]. Because of its sweetness and mouthfeel-enhancing property, erythritol has been used as a low-calorie alternative sweetener in the food industry. It is also used for prevention of dental caries, since

the bacteria that cause dental caries are unable to utilize erythritol as a carbon source [2]. In human, only about 10% of ingested erythritol is metabolized and excreted in the urine without changing insulin level [3,4]. Therefore, it has been used advantageously in special foods and pharmaceuticals for diabetic patients.

Erythritol can be synthesized from periodate-oxidized starch or dialdehyde starch by a high-temperature chemical reaction in the presence of a metal catalyst [5,6]. The chemical process has not been industrialized because it involves several intricate steps to get the erythritol and the starting material is expensive. Erythritol also can be

* Correspondence: jhseo94@snu.ac.kr

[1] Department of Agricultural Biotechnology, Seoul National University, Seoul 151-921, Korea
Full list of author information is available at the end of the article

synthesized by bioconversion that uses osmophilic yeasts and bacteria [7,8] and it has been produced commercially using an *Trichosporonoides megachiliensis* SNG-42 that produces erythritol in high yield [9]. This whole-cell bio-catalytic production process is superior to chemical sub-strate conversion because it is a simpler process requiring only a few steps and is less expensive because a raw mate-rial is cheap and readily available. Recently, a high eryth-ritol-producing yeast was isolated from honeycombs and identified as *Candida magnoliae* JH110 [10]. Various bio-logical processes using *C. magnoliae* JH110 and its chem-ical mutant have been developed for maximizing erythritol production [11-13].

Erythrose reductase (ER) catalyzes the final step of erythritol production, which is reducing erythrose to erythritol with concomitant NAD(P)H oxidation [14,15]. It has been found in yeasts and bacteria with three isozymes in a yeast strain [16]. Although there have been several reports relating to the purification and character-ization of ER from microbial sources [14,17,18], the entire primary structure and the corresponding DNA for ER still remain unknown in most of erythritol-producing yeasts. Such is the case with *C. magnoliae* JH110 [13], though one ER protein has been isolated from *C. magno-liae* KFCC-11023 [18], which is a mutant strain of *C. magnoliae* JH110 treated with ultraviolet irradiation and nitrosoguanidine for improved erythritol-producing abil-ity [7]. Indeed, the ER-encoding genes were identified only recently in *T. megachiliensis* SNG-42 [16], which is one of the yeasts used for commercial production of erythritol [9].

In this study, we provide the genetic sequence and functional characteristics of a novel NADPH-dependent ER from *C. magnoliae* JH110, which has different enzy-matic properties from the previously reported ER of *C. magnoliae* JH110 mutant strain [18]. Successful expres-sion and characterization of the functional ER could offer a potential opportunity to produce erythritol in other microorganisms, which do not produce erythritol.

Results

C. magnoliae JH110 ER gene identification

By utilizing the recently constructed expressed sequence tag (EST) library of *C. magnoliae* JH110 in our lab (unpublished data), we discovered a clone containing the ER-encoding gene from the randomly sequenced 912 EST clones. To investigate the existence of introns in this EST clone, PCR was performed using the *C. magnoliae* JH110 genomic DNA as a template. The full-sequenced PCR product had exactly the same sequence of the EST clone, indicating absence of intron. A 1041-bp *C. magnoliae* JH110 ER (CmER) with 5'- and 3'-untranslated regions was obtained from the *C. magnoliae* JH110 genomic DNA and it harbored an open reading frame (ORF) of

849 bp with an ATG initiation codon and a TGA termina-tion codon. This gene encodes a polypeptide of 282 amino acid residues with a predicted molecular mass of 31.4 kDa and an isoelectric point of 6.25. The deduced amino acid sequence of the intronless CmER gene was compared with other protein sequences of aldo-keto reductase (AKR) available from the NCBI database using the BLASTP program. The CmER showed a significant homology to the AKR superfamily. The level of identity showed the highest with aldehyde reductase I from *Aspergillus fumigatus* (XP 754700, 44% identity), and ER1 (BAD90687, 42% identity) and ER2 (BAD90688, 41% identity) from *T. megachiliensis* SNG-42. In addition, it was also highly similar to hypothetical protein sequence (XP 662433, 43% identity), which was annotated as AKR in the genome sequence of *Aspergillus nidulans*. The CmER exhibited no similarity to the well-studied xylose reductases among AKR superfamily. The multiple-sequence alignment analysis conducted using aldose reductases with erythrose reduction activity showed two conserved motifs, GYRH and AYSPL, from yeast to human (Fig. 1). The deduced amino acid sequence of the putative CmER gene was used for the construction of a phylogenetic tree with full length amino acid sequences of AKRs from various organisms [19]. In the phylogenetic tree, the CmER was close to the well-characterized glyc-erol dehydrogenase coded by *GCY1* from *Saccharomyces cerevisiae*, aldehyde reductase from *Sporobolomyces sal-monicolor*, and ERs from *T. megachiliensis* SNG-42. Fur-thermore, the CmER formed a family in the AKR clades separated from other known AKR clades. This result indicates that CmER has evolved differently from the other AKR family members (Fig. 2).

Cloning, heterologous expression, and purification

To express CmER in recombinant *Escherichia coli*, the transformant *E. coli* BL21(DE3) harboring pCmER10 was induced with 1 mM isopropyl β-D-1-thiogalactopyrano-side (IPTG) at 37°C. After IPTG induction for 6 h, a pro-tein of an apparent 33 kDa molecular weight appeared as a major band in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis (Fig. 3A). Even though an overall expression level was quite high, the protein was produced mainly as inclusion bodies. In order to express the soluble and active form of CmER in *E. coli*, the fusion gene approach was attempted. Glutathi-one S-transferase (GST) is a naturally-occurring 26 kDa protein which is often used for the soluble expression of heterologous proteins [20]. In addition, the GST fusion tag with integrated protease sites has been used for con-venient purification of the proteins of interest. Thus, the CmER fused to GST at its N-terminus was expressed under the control of the *tac* promoter by 1 mM IPTG induction in *E. coli* BL21, which also resulted in the for-
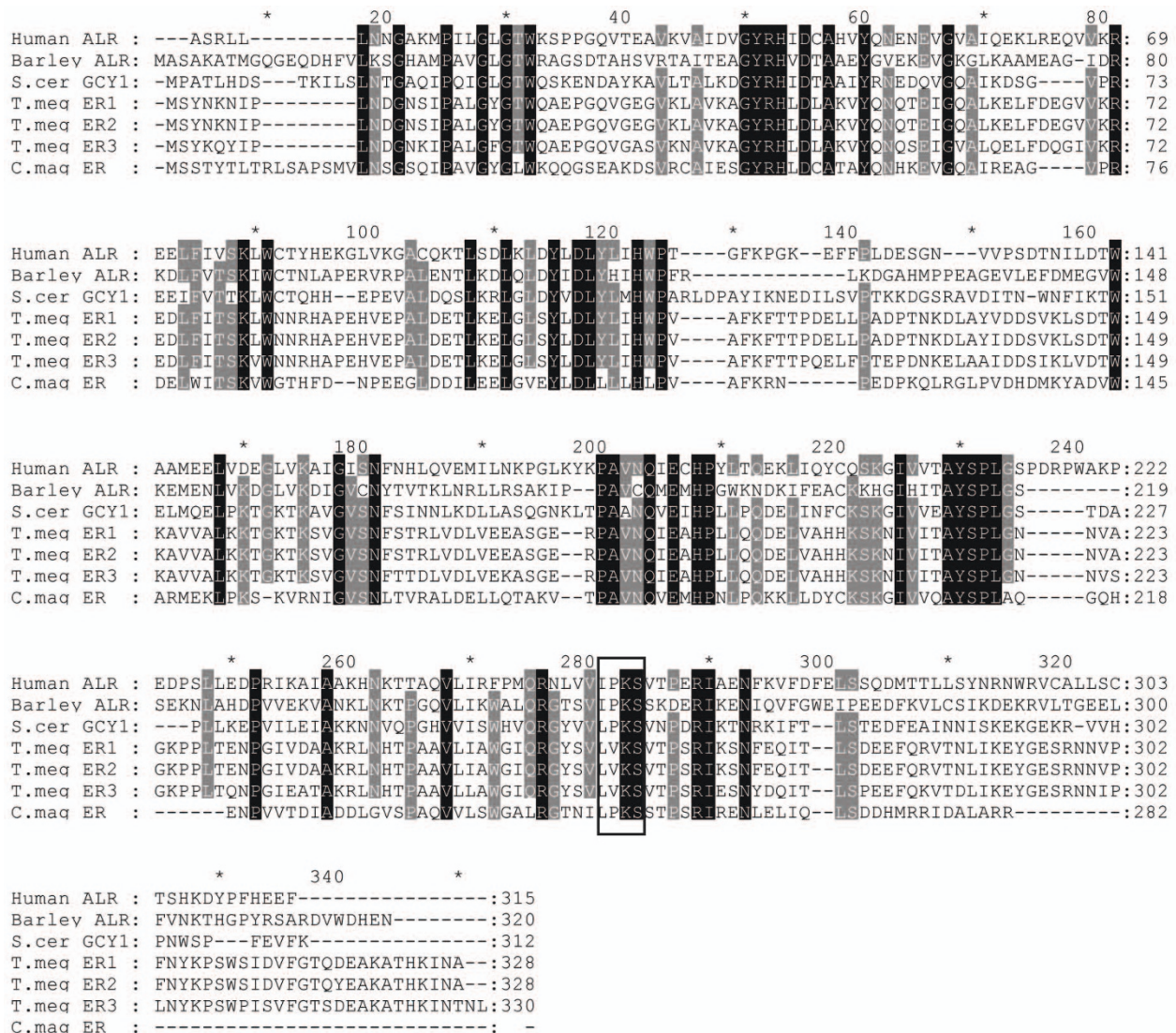
Figure 1 Multiple alignment of the deduced amino acid sequence for the ER gene from C. magnoliae JH110 with other AKRs showing erythrose reduction activity. AKRs are identified by their GenBank accession numbers: Human aldose reductase (ALR, AAB88851), barley ALR1 (CAA88322), *S. cerevisiae* GCY1 (CAA65512), *T. megachiliensis* SNG-42 ER1 (BAD90687), *T. megachiliensis* SNG-42 ER2 (BAD90688), *T. megachiliensis* SNG-42 ER3 (BAD90689), and *C. magnoliae* JH110 ER (FJ550210). The tetra-amino acid motif, IPKS, conserved among NADPH-dependent reductases is indicated by a box. Gray-shaded amino acids are conserved in at least six of the seven AKRs shown. Black-shaded amino acids are conserved in all sequences.

mation of inclusion bodies (Fig. 3B). Since it has been reported that the co-expression of molecular chaperones increases the soluble form of several enzymes in *E. coli* [21,22], their effects on CmER expression were tested in the transformant *E. coli* cells harboring pGSTCmER. Plasmid pGro7, pKJE7, or pG-KJE3 [23] was used for co-expression of GroEL-GroES, DnaK-DnaJ-GrpE, or GroEL-GroES-DnaK-DnaJ-GrpE, respectively. Plasmid pGro7, pKJE7, or pG-KJE3 was simultaneously transformed with the plasmid pGSTCmER into *E. coli* BL21 cells. Among the chaperone families co-expressed, only the GroEL-GroES-DnaK-DnaJ-GrpE chaperone family exerted positive effects on soluble expression of CmER as

verified by SDS-PAGE (Fig. 3C). After confirming soluble enzyme expression, the GST-removed CmER was purified by using the affinity column chromatography after cleavage with Factor Xa. The enzyme purity was estimated by SDS-PAGE analysis with staining of Coomassie brilliant blue. The final yield of intact CmER was 18 mg (~60 mg/liter of culture) of >95% pure CmER with a molecular mass of ~31 kDa, which corresponds well to the calculated mass (Fig. 3D).

Steady-state kinetics
Most of AKRs described to date are pyridine nucleotide linked, requiring either NADH or NADPH as a coen-
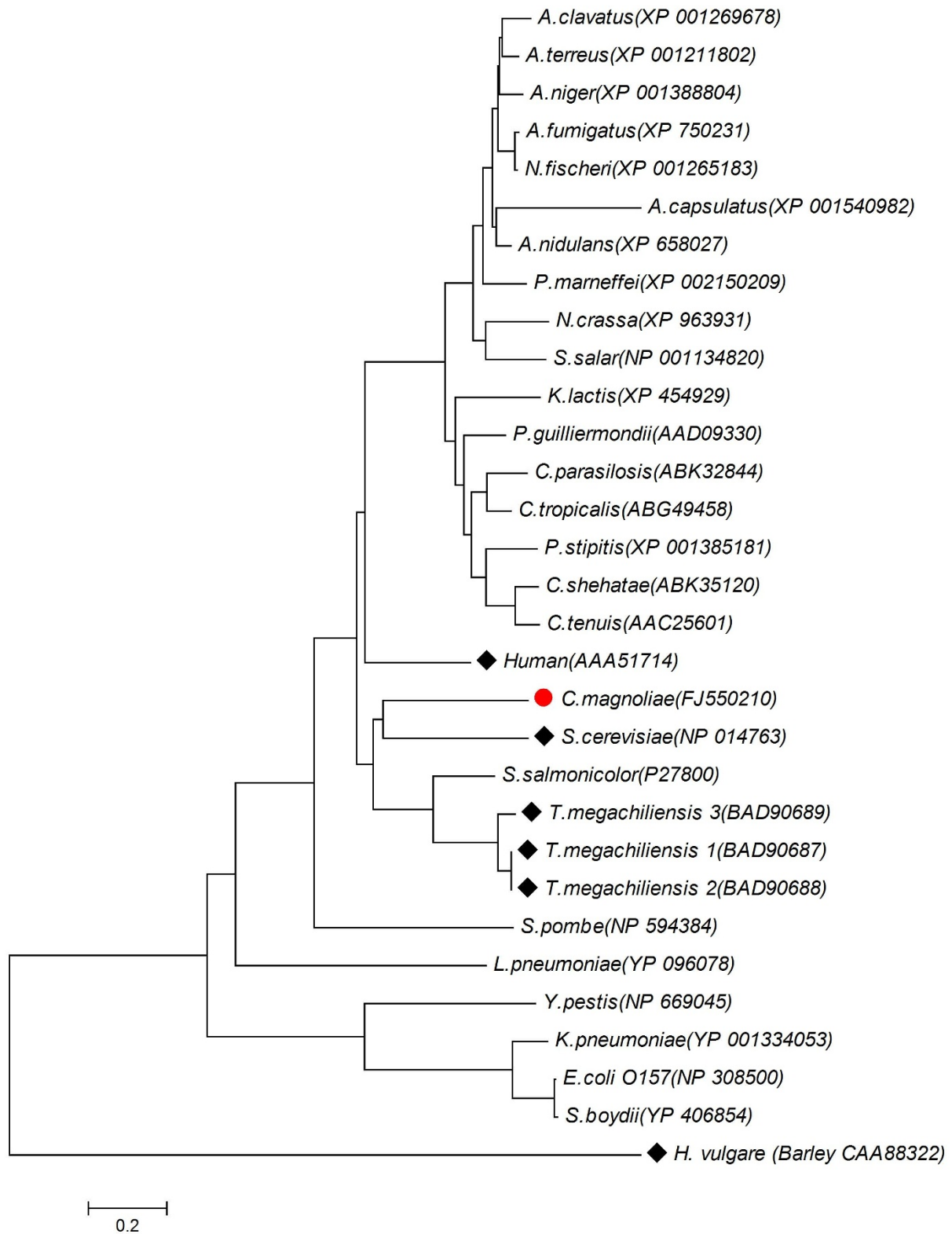
Figure 2 Phylogenetic analysis of the ER from C. magnoliae JH110. The phylogenetic tree was constructed based on the alignment of the full length amino acid sequences of AKRs including currently known erythrose reductases. AKRs are identified by their GenBank accession numbers. The AKRs used for multiple alignment analysis in Figure 1 are indicated by diamond symbols (black diamond).
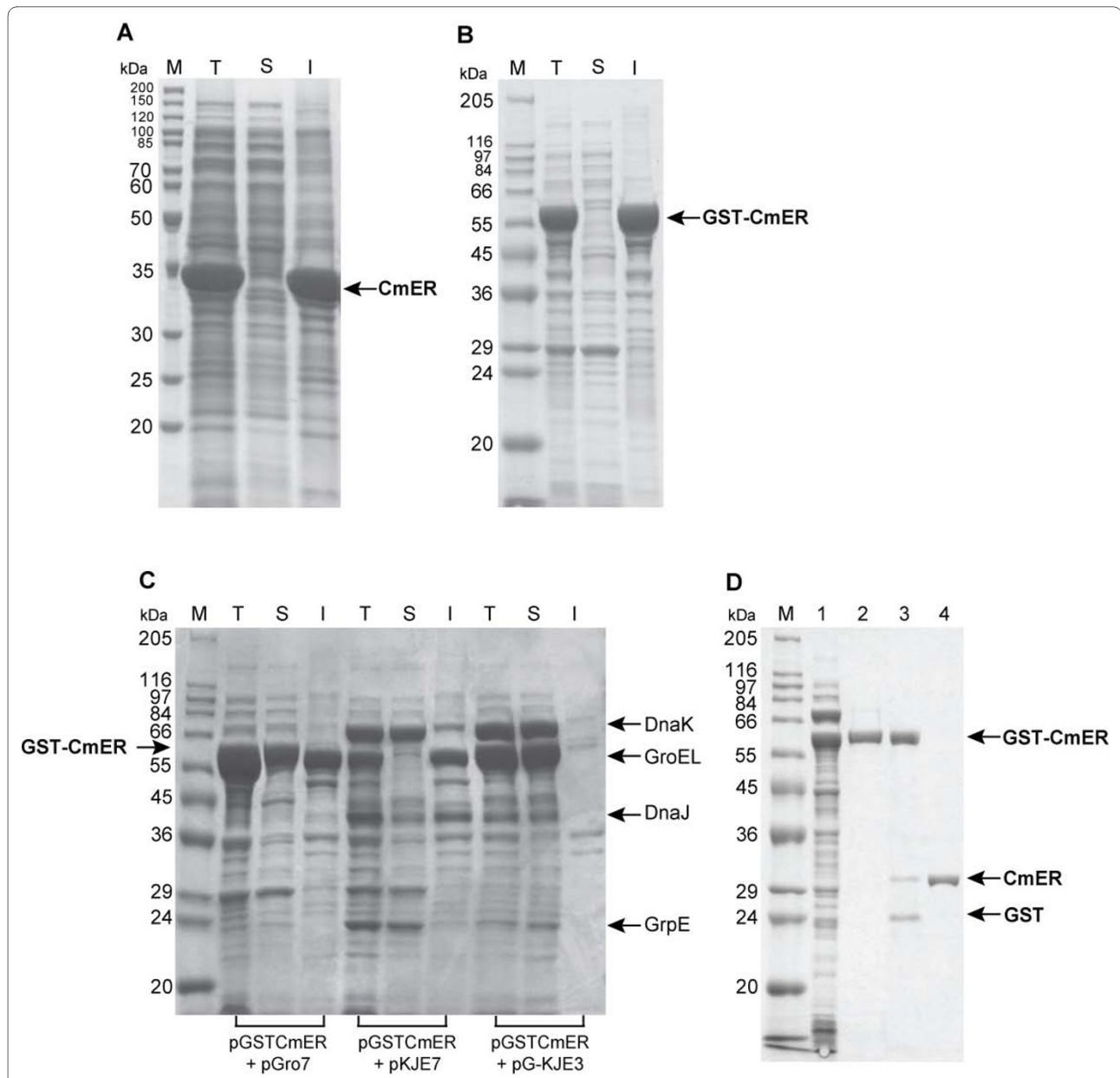
Figure 3 SDS-PAGE analysis of CmER expressed in recombinant *E. coli*. (A) Expression of CmER in the transformant *E. coli* BL21(DE3) harboring pCmER10, (B) Expressed CmER fused to GST at its N-terminus under the control of the *tac* promoter in *E. coli* BL21, (C) CmER expressed from the GST-fusion approach in conjunction with co-expression of molecular chaperones; Plasmids encoding the molecular chaperone, pGro7, pKJE7, and pGKJE3, were used to express GroEL-GroES, DnaK-DnaJ-GrpE, and GroEL-GroES-DnaK-DnaJ-GrpE, respectively, (D) SDS-PAGE analysis of CmER expressed and purified from the GST-fusion approach in conjunction with co-expression of molecular chaperones, GroEL-GroES-DnaK-DnaJ-GrpE. Lane M, protein standards; lane T, IPTG-induced total fraction; lane S, soluble fraction; lane I, insoluble fraction. Lane 1, soluble fraction of GST-fused CmER-expressed cells; lane 2, purified GST-CmER fusion protein; lane 3, cleavage of GST from GST-CmER fusion protein by Factor Xa; lane 4, purified intact CmER.

zyme. The purified CmER had activity only with NADPH (Table 1). The K_m value for NADPH was determined to be 0.016 ± 0.01 mM. In order to study the substrate specificity of the purified CmER, a series of aldoses with different numbers of carbon atoms (C4 - C6) and aromatic aldehyde *p*-nitrobenzaldehyde were tested with NADPH in excess. Kinetic parameters of the sugar substrates tested

for CmER is summarized in Table 1. As expected, the CmER displayed the highest activity toward D-erythrose in the presence of NADPH. This substrate was reduced to the corresponding alcohol, erythritol. Very low activities were found with D-xylose and D-ribose at concentration as high as 200 mM and no activity was detected with D-glucose, D-galactose, and D-fructose, which is consistent

Table 1: Steady-state kinetics for *C. magnoliae* JH110 ER

Substrate	Specific activity (U/mg protein)	K_m (mM)	k_{cat} (S^{-1})	k_{cat}/K_m (mM^{-1} S^{-1})
p-Nitrobenzaldehyde	4.3 ± 0.3	0.17 ± 0.01	1.4 ± 0.1	8.2
D-Erythrose	16.5 ± 1.3	8.5 ± 0.4	7.6 ± 0.4	0.89
D-Xylose	0.96 ± 0.06	72 ± 3.4	1.8 ± 0.1	0.011
D-Ribose	0.66 ± 0.01	56 ± 4.1	1.7 ± 0.2	0.030
D-Glucose	N.D.	N.D.	N.D.	N.D.
D-Fructose	N.D.	N.D.	N.D.	N.D.
D-Galactose	N.D.	N.D.	N.D.	N.D.
NADH	N.D.	N.D.	N.D.	N.D.
NADP$^+$	6.4 ± 0.5	0.24 ± 0.04	0.84 ± 0.06	3.5
NADPH	81.3 ± 2.7	0.016 ± 0.01	48 ± 2.5	3000

N.D.; Not Detected.

with other characterized ERs [14,24]. Moreover, the CmER shows the greatest preference for D-erythrose based on the specificity constant (k_{cat}/K_m), and the k_{cat} and k_{cat}/K_m values are considerably higher for NADPH than for NADP$^+$ (Table 1).

Temperature and pH dependence
The activity of the CmER was maximum at 42°C when assayed under standard conditions at temperatures ranging from 15 to 55°C (Fig. 4A). The activity of the CmER as a function of the pH value and the buffer system was studied between pH 4.0 and 9.0. The optimal pH was observed at 5.5 in sodium citrate buffer (Fig. 4B) and only about 20% of maximal activity was observed at pH of ≤ 4.0 and ≥ 9.0.

Homology modeling
Using the coordinates (Protein Data Bank http://www.pdb.org) for aldose reductase 1 (ALR1) from barley (PDB accession code, 2BGS) and human aldose reductase (PDB accession code, 1HQT), which have the erythrose reduction activity; a homology model docked with NADPH was created. Unlike other AKR enzymes, the resulting CmER model is an (α_8/β_6) barrel oxidoreductase with only one additional α-helix (H1). A hairpin (B1 + B2) covers one end of the barrel while a NAPDH binding site is sealing the opposite end. The H1-helix represents a junction between strand β6 and helix α7. However, the overall fold and binding of coenzyme were similar to the barley ALR1 crystal structure, as depicted in Fig. 5. The conserved catalytic residues, Trp31, Tyr60, His121, and Trp122 (CmER residue number), from barley ALR1 have similar orientations and locations in the CmER model. Structures of other AKR superfamily members display similar interactions with NADPH. In CmER, Asp54,

Asn163, Gln184, and Ser253 (CmER residue number) are the NADPH hydrogen bond partners that are completely conserved in the superfamily [25]. Furthermore, NADPH interacts with the side chains of Lys252, Thr255, and Arg258 in the CmER, which could account for the enzyme's absolute requirement of NADPH over NADH [24].

Discussion
We have isolated and characterized the CmER gene that encodes erythrose reductase from *C. magnoliae* JH110. AKRs, such as aldehyde reductase and carbonyl reductase, from *C. magnoliae* strain have been reported [26,27] and there are reports on the purification and properties of ERs from only three other yeasts [14,15,28]. However, this is the first report on the entire primary structure of the ER gene from *C. magnoliae* JH110, a microorganism with the potential for industrial erythritol production.

Initial expression of the CmER in *E. coli* resulted in formation of inclusion body. However, the efficiency of soluble CmER expression was improved using the combination of GST gene fusion and molecular chaperone co-expression (Fig. 3C). Molecular chaperones are known to play a role in protecting proteins from aggregation of unfolded or partially folded proteins in cells. It was also reported previously that co-expression of the DnaK-DnaJ-GrpE and GroEL-GroES molecular chaperone complexes improved proper folding of product proteins in *E. coli* [23]. In addition, expression of GroEL-GroES may cooperate with DnaK-DnaJ-GrpE in a synergistic way to increase soluble production of some proteins, indicating that they play cooperative roles in protein folding [23]. In agreement with this synergistic effect, soluble expression of CmER was resulted from the co-expression
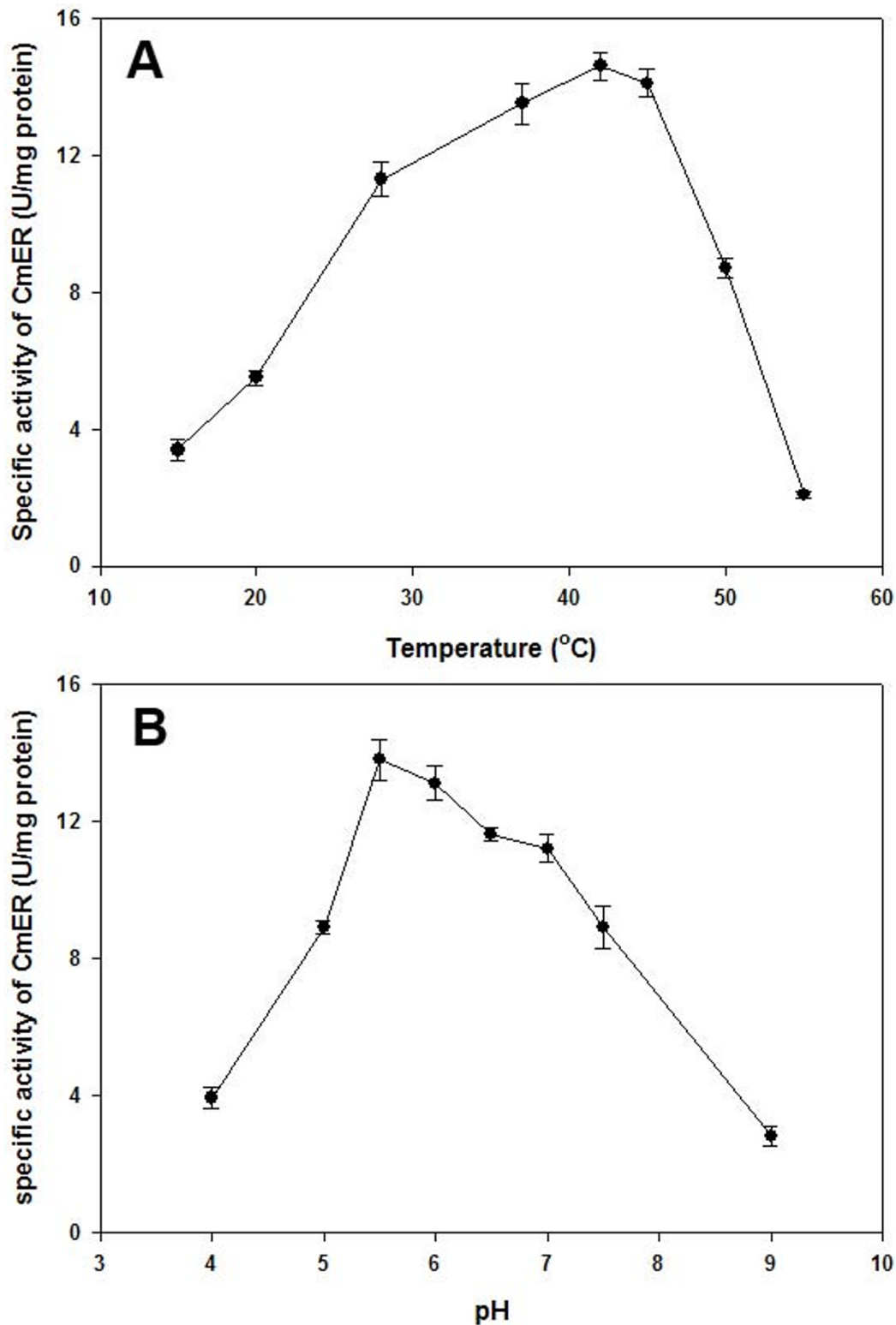
Figure 4 Determination of pH and temperature optima. (A) Effect of temperature on the specific activity of the CmER. For the effect of tempera-ture, the activity of CmER was determined at different temperatures between 15 and 55°C at pH 5.5. (B) pH dependence of CmER activity. The effect of pH on the activity of CmER was determined by assaying the enzyme activity at various pH values (pH 4.0 - 9.0). The reactions were carried out at 37°C in the following buffer systems and pH values: sodium citrate buffer (pH 4.5 - 6.0), potassium phosphate buffer (pH 6.0 - 8.0), Tris-HCl (pH 8.0 - 9.0), and glycine-NaOH buffer (pH 9.0 - 10.0). The assays were performed in triplicate at each temperature and pH point.
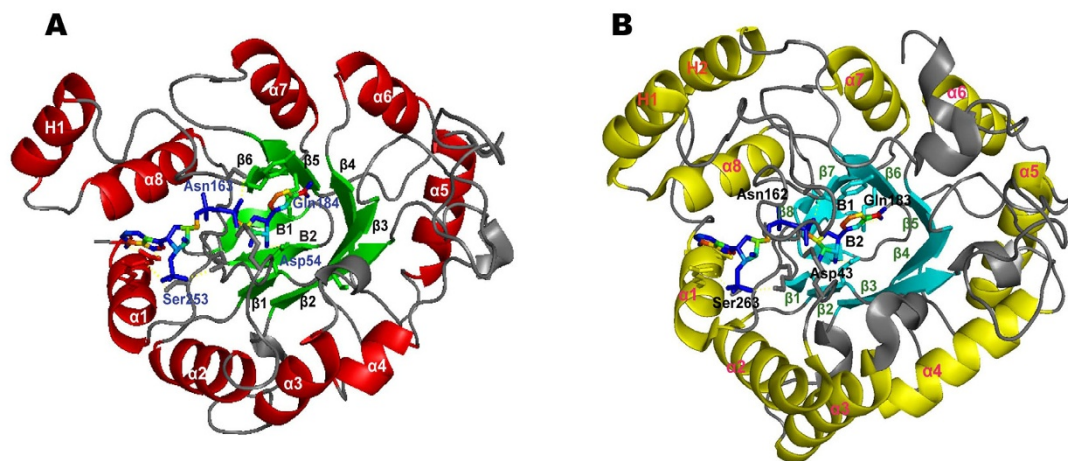
Figure 5 Comparison of the crystal structure of homology model for C. magnoliae JH110 ER (A) with barley ALR1 (B). Looking down at the NADPH cofactor in the active site at the center of the $(\alpha/\beta)_8$-barrel. A hairpin (B1 + B2) covers one end of the barrel while a NADPH binding site is sealing the opposite end. Helices are red (CmER) and yellow (barley ALR1), and sheets are green (CmER) and cyan (barley ALR1). The cofactor is included in sticks. Amino acids bound to the NADPH were colored by blue (CmER) and black (barley ALR1).

of GroEL-GroES-DnaK-DnaJ-GrpE, i.e., co-expression of single co-expression system (GroEL-GroES or DnaK-DnaJ-GrpE) was not as effective as its combined co-expression system.

Although some AKRs utilize both cofactors, NADH and NADPH [29,30], the CmER showed no activity in the presence of NADH (Table 1). This is consistent with the fact that the AKRs generally have a greater affinity for NADPH than NADH in mammalian tissues and yeasts [30,31]. Kinetic characteristics of purified and characterized ERs from *C. magnoliae* KFCC-11023, *S. cerevisiae*, *Torula corallina*, *T. megachiliensis* SNG-42, *Schizophyllum commune*, and barley are compared in Table 2. The CmER has a higher k_{cat} and higher catalytic efficiencies (k_{cat}/K_m) with respect to erythrose and NADPH. Its catalytic efficiency with respect to NADPH was 333-fold higher than that of the next closest enzyme [24]. The K_m value for D-erythrose (8.5 mM) is comparable to the values for the ERs from *T. megachiliensis* SNG-42 [15], *T. corallina* [28], and *S. commune* [32], which have K_ms of 8.2, 7.1, and 5.0 mM, respectively. In addition, the NADPH-dependent erythrose reduction was detected whereas NADP+-dependent erythritol oxidation were not observed, suggesting that under physiological conditions the formation of erythritol is thermodynamically favorable in *C. magnoliae* JH110 and hence the CmER has been better adapted to erythritol biosynthesis but not to the utilization of erythritol.

The CmER gene heterologously expressed and characterized here does not appear to encode an ER previously isolated and characterized from *C. magnoliae* KFCC-11023, which is a chemical mutant of *C. magnoliae* JH110. Due to the lack of genetic and sequence information of ER from *C. magnoliae* KFCC-11023, it is not certain to what degree they are different, but there are several key differences between the two proteins. The molecular weights and apparent kinetic constants are different (Table 1 and 2). Additionally, the previously isolated ER showed dual cofactor activity with NADH and NADPH and preferred NADH as the cofactor. The two enzymes differ in their pH optima and K_m values for D-erythrose. The enzyme, which possessed ER activity from *C. magnoliae* KFCC-11023, has been reported to have the KXXXGF(Y/G)XG motif [18]. This motif is conserved in the xylose reductase subfamily, but is not found among currently known ERs in *T. megachiliensis* SNG-42 and *C. magnoliae* JH110. Although these differences could be due to alternative mRNA splicing, posttranslational modifications in *C. magnoliae* JH110, or the fusion of the GST tag, the multiplicity of ERs has been reported in the yeast [15]. Thus, the most logical conclusion is that these are different proteins from different genes.

Barley ALR1 exhibits 15- to 30-fold higher specific activity toward short monosaccharides, for example DL-glyceraldehyde (C3) and D-erythrose (C4), than toward the osmoregulatory-relavant substrates such as xylose (C5) and glucose (C6) [24]. The barley ALR1 binds with

Table 2: Properties of ER from various organisms

ER* (reference)	Molecular mass (kDa)		Optimum		Erythrose			NADPH		
	Subunit	Native	pH	Temp. (°C)	K_m (mM)	k_{cat} (S^{-1})	k_{cat}/K_m (mM^{-1} S^{-1})	k_{cat} (S^{-1})	K_m (mM)	k_{cat}/K_m (mM^{-1} S^1)
CmER2	31	N.R.	5.5	42	8.5	7.6	0.89	48	0.016	3000
CmER1 [18]	38.8	79.0	7.0	N.R.	7.9	5.7	0.73	0.013	450	0.66
BaALR1 [24]	34	34	N.R.	N.R.	45	2.3	51	1.5	0.2	9.0
TcER [17]	35.4	71.0	6.0	40	7.12	N.R.	N.R.	N.R.	N.R.	N.R.
ScALR [32]	N.R.	N.R.	7.0	N.R.	5.0	N.R.	N.R.	N.R.	N.R.	N.R.
ScGCY1 [49]	35	N.R.	N.R.	N.R.	3.4	N.R.	N.R.	N.R.	N.R.	N.R.
TmER1 [15]	38	38	6.5	45	7.1	N.R.	N.R.	N.R.	N.R.	N.R.
TmER2 [15]	37	37	6.5	45	7.6	N.R.	N.R.	N.R.	N.R.	N.R.
TmER3 [15]	37	37	6.5	45	8.2	N.R.	N.R.	N.R.	N.R.	N.R.

*CmER2, *C. magnoliae* JH110 (This study); CmER1, *C. magnoliae* KFCC-11023; BaALR1, Barley; TcER, *T. corallina*; ScALR, *Schizophyllum commune*; ScGCY1, *S. cerevisiae*; TmER1,2, and 3, *T. megachiliensis* SNG-42. N.R.; Not Reported.

NADPH and does not react with NADH [24]. Structural models of barley ALR1 suggested that Tyr49 acts as the acid base catalyst, and Asp44, Lys79, and His112 play an important role in facilitating the hydride transfer [33,34]. All four of these amino acid residues are conserved throughout the superfamily, including the CmER. These enzymes have a strict requirement for NADPH. The tetra-amino acid motif, IPKS, is conserved among these NADPH-dependent reductases and the lysine residue in this motif is involved in NADPH binding [35]. Although the motif (LPKS) is present in the CmER, some amino acid residues were not conversed around the motif. The crystal structures of aldose reductases bound with NADPH or an analog of the cofactor show that the enzyme undergoes a large conformational change upon binding with NADPH. The change involves the reorientation of loop 7 to a position which appears to lock the coenzyme into place [36,37]. The putative loop 7 in ER from *C. magnoliae* JH110 is shorter than those in other AKRs. The K_m values for NADPH were estimated to be within the range of 2 to 4 μM for mammalian reductases [31,38] and 37.5 μM for *S. salmonicolor* reductase [39]. The difference in the K_m values could be due to an amino acid sequence difference around the NADPH-bind region. Most AKRs have a profound preference for NADPH, but remain functional with NADH [40,41]. In *Candida tenuis* XR (CtXR), which has a modest 33-fold preference for NADPH over NADH, a structural rearrangement occurs in the active site in response to the preference of either NADPH or NADH [42]. A similar rearrangement is considered impossible in CmER, not only because of its short B-loop but also due the small side chains in the B and β_6-α_8 loops (Gly216 and Ser254 in CmER versus Glu227 and Asn267 in CtXR).

Conclusions

This study provides a novel ER from *C. magnoliae* JH110, which is one of the most efficient erythritol producers. The *C. magnoliae* JH110 ER specifically acts on the erythrose with highly conserved NADPH binding site. The ease of isolating this enzyme coupled with its high activity and catalytic efficiency may be useful in the *in vitro* production of erythritol. The application value of this enzyme will be tested with *in situ* cofactor regeneration and heterologous expression in other microorganisms, which do not produce erythritol.

Methods

Chemicals and enzymes

P-Nitrobenzaldehyde, D-erythrose, D-ribose, D-xylose, D-fructose, D-galactose, D-glucose, and enzyme cofactors (NADH and NADPH) were purchased from Sigma-Aldrich (St. Louis, MO). Restriction endonucleases and other cloning reagents were purchased from New England Biolabs (Beverly, MA). The DNA samples were purified using the AccuPrep PCR purification kit (Bioneer, DaeJeon, Korea). Oligonucleotides were synthesized by Bioneer. The RediPack column packed with Glutathione Sepharose 4 Fast Flow was obtained from GE Healthcare Bio-Sciences (Uppsala, Sweden). Bacto peptone and yeast extract were from Difco Laboratories (Grayson, GA). The dye reagent for protein assay was from Bio-Rad (Hercules, CA). Unless stated otherwise, all other reagents and chemicals were from standard suppliers.

Strains and plasmids

A *C. magnoliae* JH110 wild type strain [Korean Culture Center of Microorganisms (KCCM) (formerly Korean Federation of Culture Collection) (KFCC))-10900] was maintained on YPS agar plate (10 g/L yeast extract, 20 g/L bacto peptone, and 300 g/L sucrose) at 4°C. *E. coli* DH5α was used for all DNA manipulations. *E. coli* BL21 and *E. coli* BL21(DE3) were used as expression hosts grown in Luria-Bertani (LB) medium (5 g/L yeast extract, 10 g/L tryptone, and 10 g/L NaCl) supplemented with appropriate antibiotics when required. Plasmid pETR10 containing the 10 arginine tag was derived from pET29-b(+) (Novagen, Madison, WI) [43] and used for expression of the ER gene in *E. coli* BL21(DE3). An expression vector pGEX-5X-3 (GE Healthcare Bio-Sciences) was used for expression of the ER gene as a GST-tagged fusion protein. Plasmid pGEX-5X-3 possesses a Factor Xa recognition site for cleavage of GST from the target protein. Plasmid encoding the molecular chaperone, pGro7, pKJE7, or pGKJE3 in which the L-arabinose-inducible promoter (*ara* BAD) was used to express GroEL-GroES, DnaK-DnaJ-GrpE, or GroEL-GroES-DnaK-DnaJ-GrpE, respectively, were kindly donated by Dr. Hideki Yanagi (HSP Research Institute, Kyoto Research Park, Kyoto, Japan).

Isolation of ER gene from C. magnoliae JH110

Random sequencing of the EST clones prepared from *C. magnoliae* JH110 yielded 912 successful sequencing reactions. BLAST analysis of all the ESTs revealed that an EST clone was highly similar to previously identified ER of *T. megachiliensis* SNG-42. This EST clone was named as putative erythrose reductase of *C. magnoliae* JH110 and selected for further cloning. To verify the sequence of the putative CmER gene, the gene-specific primers, ER1 (ATGTCTTCGACCTACACCCTTAC) and ER2 (TCAC-CGTCTTGCTAGCGC), were designed to amplify the sequence covering the ORF of CmER with 5' and 3' flanking regions. The PCR product was gel-purified and cloned into the pGEM-T Easy vector (Promega, Madison, WI) to facilitate DNA sequencing. After transformation into competent cells of *E. coli* DH5α, the recombinants were identified through blue-white color selection in ampicillin-containing LB-agar plates. The positive clones were sequenced in both directions and the resulting sequences were confirmed.

Sequence analysis

The searches for nucleotide and protein sequence similarity were conducted with the BLAST algorithm at the National Center for Biotechnology Information http://www.ncbi.nlm.nih.gov/blast. The deduced amino acid sequences were analyzed with the Expert Protein Analysis System http://www.expasy.org/ and protein domain features of CmER were determined by using the Simple Modular Architecture Research Tool [44]. The deduced amino acid sequence of CmER was aligned with the corresponding AKR sequences including ER from various organisms using the CLUSTALW software [45]. A phylogenetic tree was constructed with MEGA 3.1 software by using the neighbor-joining method [46]. Boostrap analysis was used with 1000 replicates to test the relative support for the branches produced by the neighbor-joining analysis [47]. All the analyzed sequences of aldose reductase enzymes were retrieved from GenBank and SWISS-PROT database.

Expression and purification

After identifying the EST sequence of CmER, the CmER gene was amplified from the *C. magnoliae* JH110 genomic DNA by PCR using the primer ER3 (GGAAT-TC**CATATG**TCTTCGACCTACACC, NdeI restriction site is highlighted as bold) and ER4 (CGC**GGATC-C**CCGTCTTGCTAGCGCG, BamHI restriction site is highlighted as bold). The amplified CmER gene was sequenced and compared with the sequence of the identified putative CmER EST clone to investigate the existence of introns. For expression of the CmER gene, the amplified CmER was cloned into plasmid pETR10, resulting in pCmER10. For fusion gene approach, the amplified CmER with a set of primers (ER5; CGC**GGATC-C**CCTCTTCGAACCTACACCCTTACTCG, BamHI restriction site is highlighted as bold and ER6; ATAA-GAAT**GCGGCCGC**TCACCGTCTTGCTAGCGC, NotI restriction site is highlighted as bold) was inserted into plasmid pGEX-5X-3 to express the GST-fused CmER, resulting in pGSTCmER. The constructed recombinant vectors, pCmER10 and pGSTCmER, were introduced into *E. coli* BL21 (DE3) and *E. coli* BL21 cells, respectively. Transformant cells were selected on LB-agar plates containing appropriate antibiotics and cultured in LB media at 37°C with vigorous shaking. Expression of the CmER and molecular chaperone genes were induced at the logarithmic phase of growth ($OD_{600} \approx 0.5$) by adding IPTG and L-arabinose to a final concentration of 1 mM and 1 g/L, respectively. The cells induced for 6 h were then harvested by centrifugation at $6,000 \times g$ for 20 min at 4°C and resuspended in $1 \times$ PBS buffer (pH 7.4) containing the protease inhibitor cocktail (Sigma-Aldrich) for disruption by sonication. The crude extract was fractionated into soluble and insoluble fractions by centrifugation at $20,000 \times g$ for 30 min at 4°C. These fractions were analyzed by SDS-PAGE and the soluble fraction was used for the subsequent purification of the CmER enzyme. ÄKTA FPLC (GE Healthcare Bio-Sciences), an automated chromatography system, was used to analyze the separation pattern of the fusion protein. The soluble fraction was directly applied to the GSTrap FF column (1 mL, GE Healthcare Bio-Sciences) prepacked with Glutathione Sepharose 4 Fast Flow (GE Healthcare Bio-Sciences). The column was washed with $1 \times$ PBS buffer (pH 7.4) at a flow rate of 0.5 mL/min. After washing, on-column cleavage was per-

formed using Factor Xa dissolved in cleavage buffer (1 mM $CaCl_2$ and 150 mM NaCl in 50 mM Tris-HCl, pH 7.5) and then the bound protein was eluted. For removal of Factor Xa protease, the eluate was passed directly into the column manually packed with Benzamidine 4 Fast Flow (GE Healthcare Bio-Sciences). The final eluate without Factor Xa was analyzed for the SDS-PAGE and enzymatic assay.

Determination of protein content and enzyme activity

Protein content was determined using a protein assay kit (Bio-Rad) with bovine serum albumin as standard. The CmER activity was determined as reported previously [14] with modifications given below. Five microliter of the purified enzyme solution was incubated in 250 µL of 50 mM phosphate buffer (pH 6.5) containing 12 mM D-erythrose and 4 mM NADH or NADPH at 37°C for 10 min. Immediately after the reaction, the absorbance at 340 nm was measured by an Ultrospec 4000 spectrophotometer (GE Healthcare Bio-Sciences). One unit of enzyme activity was defined as the amount of enzyme which produced 1 µmol of $NAD(P)^+$ per minute under the above conditions.

Effects of pH and temperature

The influence of pH on the activity of CmER was studied by assaying the enzyme activity at various pH values (pH 4.0 - 9.0). The reactions were carried out at 37°C in the following buffer systems and pH values: sodium citrate buffer (pH 4.5 - 6.0), potassium phosphate buffer (pH 6.0 - 8.0), Tris-HCl (pH 8.0 - 9.0), and glycine-NaOH buffer (pH 9.0 - 10.0). D-Erythrose was a variable substrate and NADPH remained at a fixed concentration of 50 µM. For the effect of temperature, the activity of CmER was determined at different temperatures between 15 and 55°C at pH 5.5. The assays were performed in triplicate at each temperature and pH point.

Steady-state kinetics

Kinetic parameters of purified CmER were determined by steady-state kinetics experiments. The enzyme reaction for D-erythrose reduction was performed at optimal pH 5.5 and temperature 42°C by varying concentrations of one substrate (D-erythrose or NADPH) while the other was kept constant under saturated conditions. The kinetic parameters, K_m and V_{max}, were determined by plotting the reaction rates against substrate concentrations to fit the Michaelis-Menten kinetics.

SDS-PAGE analysis

SDS-PAGE was performed according to Laemmli [48] using a Mini PROTEAN 3 electrophoresis system (Bio-Rad). The gels were visualized by staining with Coomassie brilliant blue R-250.

Homology modeling

Molecular Operating Environment (MOE, Chemical Computing Group Inc., Montreal, Canada) was used to prepare and optimize the model and Insight II (version 2000; Accelrys Inc., San Diego, CA) was used for analysis. To verify the model, the overall fold was checked using Profiles-3D (Insight II), and the allowed states for φ and ψ angles and bond distances were checked using ProStat (Insight II), both with default settings. The Profiles-3D check resulted in a self-compatibility score of 86.5, which compares well to the scores of 98.3, 93.5 and 106.2 for the coordinates from 1ADS, 1BGS and 1HQT, respectively. The ProStat check of φ and ψ angles were determined to be 75.2% within the allowed Ramanchandran region, comparing well to the 87.1%, 89.4 and 84.6% for the same analysis of PDB structures 1ADS, 1BGS, and 1HQT, respectively and lying in the core region of the Ramachandran plot. Therefore, our calculated results imply that the modeled structure of CmER is reliable.

Nucleotide sequence accession number

The nucleotide sequence of genomic CmER gene has been submitted to the GenBank database under accession number FJ550210.

Author Details

[1]Department of Agricultural Biotechnology, Seoul National University, Seoul 151-921, Korea, [2]Department of Molecular Science and Technology, Ajou University, Suwon 442-749, Korea and [3]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA

References

1. Goossens J, Röper H: **Erythritol: a new sweetener.** *Food Sci Tech Today* 1994, **8**:144-148.
2. Munro IC, Bernt WO, Borzelleca JF, Flamm G, Lynch BS, Kennepohl E, Bar EA, Modderman J: **Erythritol: an interpretive summary of biochemical, metabolic, toxicological and clinical data.** *Food Chem Toxicol* 1998, **36**(12):1139-1174.
3. Hiele M, Ghoos Y, Rutgeerts P, Vantrappen G: **Metabolism of erythritol in humans: comparison with glucose and lactitol.** *Br J Nutr* 1993, **69**(1):169-176.
4. Arrigoni E, Brouns F, Amado R: **Human gut microbiota does not ferment erythritol.** *Br J Nutr* 2005, **94**(5):643-646.
5. Pfeifer VF, Sohns VE, Conway HF, Lancaster EB, Dabic S, Griffin EL: **Two stage process for dialdehyde starch using electrolytic regeneration of periodic acid.** *Ind Eng Chem* 1960, **52**(3):201-206.
6. Otey FH, Sloan JW, Wilhan CA, Mehltretter CL: **Erythritol and ethylene glycol from dialdehyde starch.** *Ind Eng Chem* 1961, **53**:267-268.
7. Ryu YW, Park CY, Park JB, Kim SY, Seo JH: **Optimization of erythritol production by *Candida magnoliae* in fed-batch culture.** *J Ind Microbiol Biotechnol* 2000, **25**(2):100-103.
8. Veiga-da-Cunha M, Santos H, Van Schaftingen E: **Pathway and regulation of erythritol formation in Leuconostoc oenos.** *J Bacteriol* 1993, **175**(13):3941-3948.
9. Ishizuka H, Wako K, Kasumi T, Sasaki T: **Breeding of a mutant of *Aureobasidium* sp. with high erythritol production.** *J Ferment Bioeng* 1989, **68**(5):310-314.

10. Kim SY, Park SS, Jeon YJ, Seo JH: **Analysis of fermentation characteristics for production of erythritol by *Candida* sp.** *Kor J Food Sci Technol* 1996, **28(5)**:935-939.

11. Park SY, Seo JH, Ryu YW: **Two-stage fed-batch culture of *Candida magnoliae* for the production of erythritol using an industrial medium.** *Kor J Biotechnol Bioeng* 2003, **18(4)**:249-254.

12. Yang SW, Park JB, Han NS, Ryu YW, Seo JH: **Production of erythritol from glucose by an osmophilic mutant of *Candida magnoliae*.** *Biotechnol Lett* 1999, **21**:887-890.

13. Koh ES, Leet TH, Lee DY, Kim HJ, Ryu YW, Seo JH: **Scale-up of erythritol production by an osmophilic mutant of *Candida magnoliae*.** *Biotechnol Lett* 2003, **25(24)**:2103-2105.

14. Isuizuka H, Tokuoka K, Sasaki T, Taniguchi H: **Purification and some properties of an erythrose reductase from an *Aureobasidium* sp. mutant.** *Biosci Biotechnol Biochem* 1992, **56(6)**:941-945.

15. Tokuoka K, Ishizuka H, Wako K, Taniguchi H: **Comparison of three-forms of erythrose reductase from an *Aureobasidium* sp. mutant.** *J Gen Appl Microbiol* 1992, **38(22)**:145-155.

16. Ookura T, Azuma K, Isshiki K, Taniguchi H, Kasumi T, Kawamura Y: **Primary structure analysis and functional expression of erythrose reductases from erythritol-producing fungi (*Trichosporonoides megachiliensis* SNG-42).** *Biosci Biotechnol Biochem* 2005, **69(5)**:944-951.

17. Lee JK, Hong KW, Kim SY: **Purification and properties of a NADPH-dependent erythrose reductase from the newly isolated *Troula corallina*.** *Biotechnol Prog* 2003, **19**:495-500.

18. Lee JK, Kim SY, Ryu YW, Seo JH, Kim JH: **Purification and characterization of a novel erythrose reductase from *Candida magnoliae*.** *Appl Environ Microbiol* 2003, **69(7)**:3710-3718.

19. Saitou N, Nei M: **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Mol Biol Evol* 1987, **4(4)**:406-425.

20. Wu X, Oppermann U: **High-level expression and rapid purification of rare-codon genes from hyperthermophilic archaea by the GST gene fusion system.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2003, **786(1-2)**:177-185.

21. Lee DH, Kim MD, Lee WH, Kweon DH, Seo JH: **Consortium of fold-catalyzing proteins increases soluble expression of cyclohexanone monooxygenase in recombinant *Escherichia coli*.** *Appl Microbiol Biotechnol* 2004, **63(5)**:549-552.

22. Kim SG, Kweon DH, Lee DH, Park YC, Seo JH: **Coexpression of folding accessory proteins for production of active cyclodextrin glycosyltransferase of *Bacillus macerans* in recombinant *Escherichia coli*.** *Protein Expres Pur* 2005, **41(2)**:426-432.

23. Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T: **Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*.** *Appl Environ Microbiol* 1998, **64(5)**:1694-1699.

24. Renza R, Francesco S, Dorothea B: **An aldose reductase homologous gene from barley: regulation and function.** *Plant J* 1995, **7(5)**:809-822.

25. Jez JM, Bennett MJ, Schlegel BP, Lewis M, Penning TM: **Comparative anatomy of the aldo-keto reductase superfamily.** *Biochem J* 1997, **326(3)**:625-636.

26. Wada M, Kataoka M, Kawabata H, Yasohara Y, Kizaki N, Hasegawa J, Shimizu S: **Purification and characterization of NADPH-dependent carbonyl reductase, involved in stereoselective reduction of ethyl 4-chloro-3-oxobutanoate, from *Candida magnoliae*.** *Biosci Biotechnol Biochem* 1998, **62(2)**:280-285.

27. Wada M, Kawabata H, Kataoka M, Yasohara Y, Kizaki N, Hasegawa J, Shimizu S: **Purification and characterization of an aldehyde reductase from *Candida magnoliae*.** *J Mol Catal B: Enz* 1999, **6(3)**:333-339.

28. Lee JK, Koo BS, Kim SY: **Fumarate-mediated inhibition of erythrose reductase, a key enzyme for erythritol production by *Torula corallina*.** *Appl Environ Microbiol* 2002, **68(9)**:4534-4538.

29. Neuhauser W, Haltrich D, Kulbe K, Nidetzky B: **NAD(P)H-dependent aldose reductase from the xylose-assimilating yeast *Candida tenuis*. Isolation, characterization and biochemical properties of the enzyme.** *Biochem J* 1997, **326(Pt 3)**:683-692.

30. Verduyn C, Van Kleef R, Frank J, Schreuder H, Van Dijken J, Scheffers W: **Properties of the NAD(P)H-dependent xylose reductase from the xylose-fermenting yeast *Pichia stipitis*.** *Biochem J* 1985, **226(3)**:669-677.

31. Bendicht W, Hanspeter BÜR, Kurt B, Jean-Pierre W: **Purification and characterization of human-brain aldose reductase.** *Eur J Biochem* 1982, **127(2)**:279-284.

32. Braun ML, Niederpruem DJ: **Erythritol metabolism in wild-type and mutant strains of *Schizophyllum commune*.** *J Bacteriol* 1969, **100(2)**:625-634.

33. Bohren K, Grimshaw C, Lai C, Harrison D, Ringe D, Petsko G, Gabbay K: **Tyrosine-48 is the proton donor and histidine-110 directs substrate stereochemical selectivity in the reduction reaction of human aldose reductase: enzyme kinetics and crystal structure of the Y48H mutant enzyme.** *Biochemistry* 1994, **33(8)**:2021-2032.

34. Tarle I, Borhani DW, Wilson DK, Quiocho FA, Petrash JM: **Probing the active site of human aldose reductase. Site-directed mutagenesis of Asp-43, Tyr-48, Lys-77, and His-110.** *J Biol Chem* 1993, **268(34)**:25687-25693.

35. Bohren KM, Page JL, Shankar R, Henry SP, Gabbay KH: **Expression of human aldose and aldehyde reductases. Site-directed mutagenesis of a critical lysine 262.** *J Biol Chem* 1991, **266(35)**:24031-24037.

36. Borhani DW, Harter TM, Petrash JM: **The crystal structure of the aldose reductase. NADPH binary complex.** *J Biol Chem* 1992, **267(34)**:24841-24847.

37. Rondeau J, Tete-Favier F, Podjarny A, Reymann J, Barth P, Biellmann J, Moras D: **Novel NADPH-binding domain revealed by the crystal structure of aldose reductase.** *Nature* 1992, **355(6359)**:469-472.

38. Wermuth B, Munch J, von Wartburg J: **Purification and properties of NADPH-dependent aldehyde reductase from human liver.** *J Biol Chem* 1977, **252(11)**:3821-3828.

39. Kataoka M, Sakai H, Morikawa T, Katoh M, Miyoshi T, Shimizu S, Yamada H: **Characterization of aldehyde reductase of *Sporobolomyces salmonicolor*.** *Biochim Biophys Acta* 1992, **1122(1)**:57-62.

40. Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S: **Alteration of the specificity of the cofactor-binding pocket of *Corynebacterium* 2,5-diketo-D-gluconic acid reductase A.** *Protein Eng* 2002, **15(2)**:131-140.

41. Kubiseski TJ, Flynn TG: **Studies on human aldose reductase.** *J Biol Chem* 1995, **270(28)**:16911-16917.

42. Petschacher B, Leitgeb S, Kavanagh KL, Wilson DK, B N: **The coenzyme specificity of *Candida tenuis* xylose reductase (AKR2B5) explored by site-directed mutagenesis and X-ray crystallography.** *Biochem J* 2005, **385(Pt 1)**:75-83.

43. Lee DH, Kim SG, Kweon DH, Seo JH: **Folding machineries displayed on a cation-exchanger for the concerted refolding of cysteine- or proline-rich proteins.** *BMC Biotechnol* 2009, **9(1)**:27-36.

44. Schultz J, Milpetz F, Bork P, Ponting CP: **SMART, a simple modular architecture research tool: Identification of signaling domains.** *Proc Natl Acad Sci* 1998, **95(11)**:5857-5864.

45. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD: **Multiple sequence alignment with the clustal series of programs.** *Nucleic Acids Res* 2003, **31(13)**:3497-3500.

46. Kumar S, Tamura K, Nei M: **MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment.** *Brief Bioinform* 2004, **5(2)**:150-163.

47. Felsenstein J: **Taking variation of evolutionary rates between sites into account in inferring phylogenies.** *J Mol Evol* 2001, **53(4-5)**:447-455.

48. Laemmli UK: **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature* 1970, **227(5259)**:680-685.

49. Ookura T, Kasumi T: **Yeast Gcy1p reduces erythrose and erythrose-4-phosphate.** *Rep Natl Food Res Inst* 2007, **71**:57-60.